Supplement for "Estimation of long range dependence in gappy Gaussian time series"

Peter F. Craigmile[1] and **Debashis Mondal**[2]

[1] Department of Statistics, The Ohio State University, Columbus, OH 43210, USA.
E-mail: pfc@stat.osu.edu

[2] Department of Statistics, Oregon State University, Corvallis, OR 97331, USA.
E-mail: debashis@stat.oregonstate.edu

Supplemental figures and tables

Figures S1 and S2 are presented in addition to Figure 2 in the text. Figure S1 sets $N = 1024$, $\delta = 0.40$ with a D(4) wavelet filter, and varies the choice of the levels J_0 and J. Based on 10,000 simulations of FD processes, we compare the bias, SD, and RMSE of the full and diagonal estimators. Figure S2 sets $N = 2048$. In both figures, each row shows a different value of J_0 ($J_0 = 1, 2, 3, 4$) and each column plots the bias, SD, or RMSE versus J. The diagonal estimator is denoted by the gray circles, and the full estimator is denoted by the black triangles. 95% bootstrap confidence intervals for each statistic are smaller in height than the height of the symbols, and thus are omitted. The patterns are similar to the $N = 8192$ case discussed in the main article.

Tables S1 and S2 compare the bias and estimated standard errors for the diagonal and full estimators for an FD(δ) when $\delta = 0.25$ and $\delta = 0.80$, respectively. These tables are presented in addition to Table 2 in the main article.

Figure S3 complements Figure 5 in the main text, and summarizes the log wavelet variance of the deseasonalized and standardized temperature series versus the log wavelet scale, at all 17 locations.
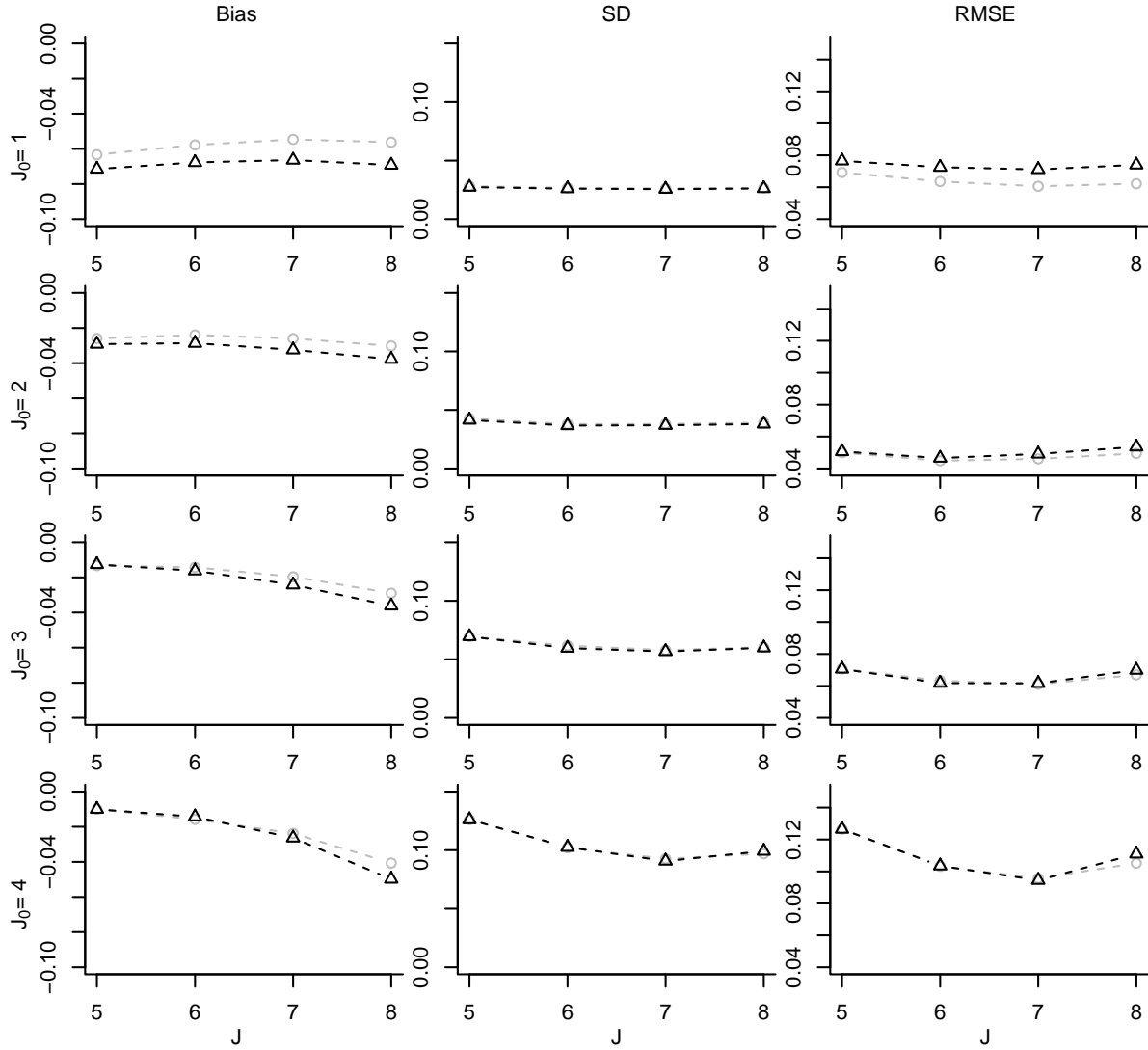
Figure S1: When $N = 1024$ and $\delta = 0.40$, a comparison of the bias, SD, and RMSE for the diagonal (gray circles) and full (black triangles) estimator versus J. Each row shows a different value for J_0.
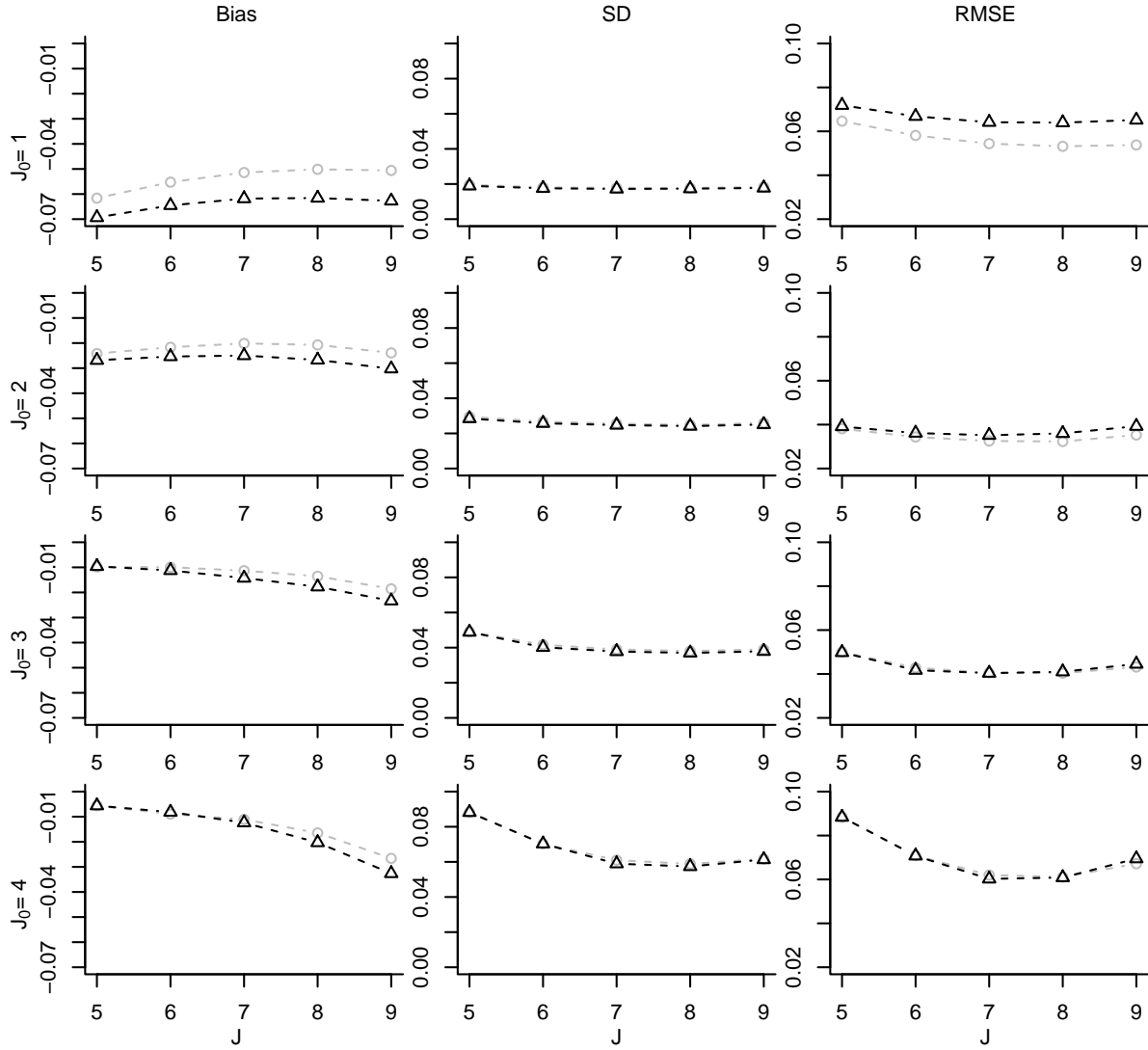
Figure S2: When $N = 2048$ and $\delta = 0.40$, a comparison of the bias, SD, and RMSE for the diagonal (gray circles) and full (black triangles) estimator versus J. Each row shows a different value for J_0.

3

Table S1: Comparisons of the bias and estimated standard errors, for the diagonal and full estimators of the LRD parameter for an FD(0.25) process, under a different number of non-missing and missing data senarios. In each case the D(4) wavelet filter is used with $J_0 = 3$, and $J = 7$. The bootstrap SE for the bias is no larger than 0.0008; the bootstrap SE for the SD is no larger than 0.0005; the SE for the averaged ESE is no larger than 3×10^{-5}.

	No gaps			Binomial, 10% gaps	Markov
	Cov	$R = 7$	$R = 9$	$R = 7$	$R = 7$
Diagonal, Bias	-0.0177	-0.0160	-0.0155	-0.0160	-0.0159
Full, Bias	-0.0217	-0.0186	-0.0174	-0.0169	-0.0174
Diagonal, SD	0.0574	0.0612	0.0613	0.0638	0.0614
Diagonal, Avg. ESE	0.0434	0.0343	0.0346	0.0366	0.0354
Diagonal, Avg. Sandwich ESE	0.0513	0.0378	0.0377	0.0401	0.0388
Full, SD	0.0564	0.0868	0.0722	0.0895	0.0862
Full, Avg. ESE	0.0490	0.0268	0.0299	0.0283	0.0274

Table S2: Comparisons of the bias and estimated standard errors, for the diagonal and full estimators of the LRD parameter for an FD(0.80) process, under a different number of non-missing and missing data senarios. In each case the D(4) wavelet filter is used with $J_0 = 3$, and $J = 7$. The bootstrap SE for the bias is no larger than 0.0008; the bootstrap SE for the SD is no larger than 0.0005; the SE for the averaged ESE is no larger than 3×10^{-5}.

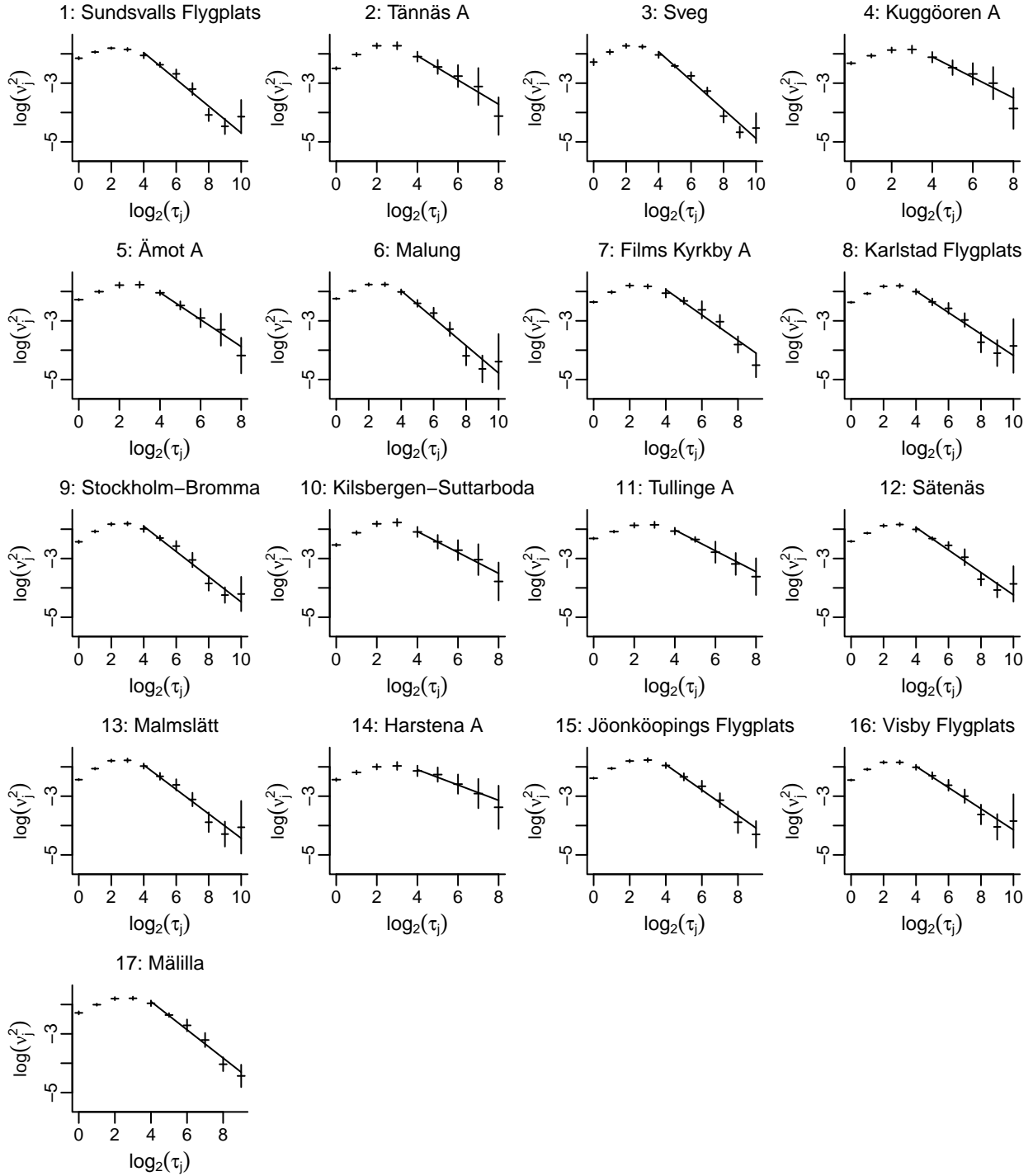|	No gaps			Binomial, 10% gaps	Markov
	Cov	$R = 7$	$R = 9$	$R = 7$	$R = 7$
Diagonal, Bias	-0.0217	-0.0200	-0.0199	-0.0187	-0.0195
Full, Bias	-0.0273	-0.0226	-0.0215	-0.0209	-0.0208
Diagonal, SD	0.0610	0.0650	0.0647	0.0665	0.0656
Diagonal, Avg. ESE	0.0448	0.0355	0.0362	0.0368	0.0361
Diagonal, Avg. Sandwich ESE	0.0544	0.0400	0.0401	0.0414	0.0406
Full, SD	0.0593	0.0907	0.0748	0.0904	0.0884
Full, Avg. ESE	0.0508	0.0277	0.0313	0.0286	0.0282

Figure S3: The log wavelet variance of the deseasonalized and standardized temperature series versus the log wavelet scale, at all 17 locations. The horizontal line indicates the estimates, and the vertical bars gives their 95% confidence intervals. The straight line across time scales is based on the diagonal estimator of the LRD parameter.